

W# 5000



PCT10

## RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/10/069,425A

TIME: 12:30:18

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\06252003\J069425A.raw

3 <110> APPLICANT: E.I. du Pont de Nemours and Company, Inc.  
 6 <120> TITLE OF INVENTION: cis-Prenyltransferases from Plants  
 8 <130> FILE REFERENCE: BC1019 PCT  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/069,425A  
 C--> 10 <141> CURRENT FILING DATE: 2002-02-19  
 10 <150> PRIOR APPLICATION NUMBER: PCT/US 00/25856  
 11 <151> PRIOR FILING DATE: 2000-09-21  
 13 <150> PRIOR APPLICATION NUMBER: US 60/155,046  
 14 <151> PRIOR FILING DATE: 1999-09-21  
 16 <160> NUMBER OF SEQ ID NOS: 37  
 18 <170> SOFTWARE: PatentIn version 3.1  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 1388  
 22 <212> TYPE: DNA  
 23 <213> ORGANISM: Dimorphotheca  
 25 <400> SEQUENCE: 1

26 ggcacgacag gtttcccgac tggaaagcgg gcagtgcgag caacgcaatt aatgtgagtt 60  
 28 agctcactca ttaggcaccg caggctttac actttatgct tccggctcgt atgttggtgtg 120  
 30 gaattgtgag cggataacaa ttccacacag gaaacagcta tgaccatgat tacgccaagc 180  
 32 gcgcaattaa ccctcactaa agggaacaaa aggctggagc tccaccgcgg tggcgggccgc 240  
 34 tctagaacta gtggatcccc cgggctgcag gaattcggca cgagcttaaa taatgcttaa 300  
 36 tcttccccctc tacttaccce aatatccctg ttatttcccg gcctctctct ccaccaacca 360  
 38 ccaccgtggt ctttatgtat tcaaccaatc agacaccact ggaggtggaa ttaattcgct 420  
 40 ggaggaacgc attactccag caggactcaa gcacgagtta atgccaaagc atgtggcagt 480  
 42 gatcatggat ggaacacagga gatgggctcg atcacgtggg ttaatgccgg atgctgggta 540  
 44 catggaaggt gcacgctcat tgaaggtgat ggtggaattg tgcgtaaat ggggaattca 600  
 46 agtccttact gtgtttgct tctcagctga taactgggta agacccaaag ttgaagttga 660  
 48 tttcttgatg ggactaattg aaagtgtatt aaaagatgaa gttgttcata tgatcaaaga 720  
 50 gggatatccag ctttcgggta tcggagacac atctaagctt ccaaaatcgg taaaacggat 780  
 52 cattacatat gctgaaaaca tcacgaagaa caactcaca ctcaatcttg ttgtagcaat 840  
 54 aaattatagt ggaataatag atatcgtcca agcttgtcaa agcatcgac taaaagtcaa 900  
 56 agacggtgtc attcaacccg aagaaatcaa tgagtttacg attgaaaatg aacttggtac 960  
 58 aaattgtatt ccttttccac accctgatct actaattcgg actagtggg agcttagagt 1020  
 60 gagcaacttc tttttgtggc aattggcgta tactgaatta tacttcagt aaactctttg 1080  
 62 gcctgatttt ggtgaagatg aacttttaca tgccttaaat acttttcaac atagacgaag 1140  
 64 acgttatggt ggatgagatt cttaaacac cctgtagagt tgcataatcat attgactttt 1200  
 66 gatattgttt aatactattt atattattat tatgttgtaa tctcgtacta gaacatgaat 1260  
 68 ttaaataaggc aatagagcat gccaccta atgtctagtt atgagattct aaagacgtaa 1320  
 70 ttatgcttac ctaaaagaaa atatatatga agagaaaagc ttatgtaaaa aaaaaaaaaa 1380  
 72 aaaaaaaaaa 1388  
 75 <210> SEQ ID NO: 2  
 76 <211> LENGTH: 287  
 77 <212> TYPE: PRT

p. 6  
ENTERED

## RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/10/069,425A

TIME: 12:30:18

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\06252003\J069425A.raw

```

78 <213> ORGANISM: Dimorphotheca
80 <400> SEQUENCE: 2
82 Met Leu Asn Leu Pro Leu Tyr Leu Pro Lys Tyr Pro Cys Tyr Phe Pro
83 1 5 10 15
86 Ala Ser Leu Ser Thr Asn His His Arg Gly Leu Tyr Val Phe Asn Gln
87 20 25 30
90 Ser Asp Thr Thr Gly Gly Gly Ile Asn Ser Leu Glu Glu Arg Ile Thr
91 35 40 45
94 Pro Ala Gly Leu Lys His Glu Leu Met Pro Lys His Val Ala Val Ile
95 50 55 60
98 Met Asp Gly Asn Arg Arg Trp Ala Arg Ser Arg Gly Leu Met Pro Asp
99 65 70 75 80
102 Ala Gly Tyr Met Glu Gly Ala Arg Ser Leu Lys Val Met Val Glu Leu
103 85 90 95
106 Cys Arg Lys Trp Gly Ile Gln Val Leu Thr Val Phe Ala Phe Ser Ala
107 100 105 110
110 Asp Asn Trp Leu Arg Pro Lys Val Glu Val Asp Phe Leu Met Gly Leu
111 115 120 125
114 Ile Glu Ser Val Leu Lys Asp Glu Val Val His Met Ile Lys Glu Gly
115 130 135 140
118 Ile Gln Leu Ser Val Ile Gly Asp Thr Ser Lys Leu Pro Lys Ser Val
119 145 150 155 160
122 Lys Arg Ile Ile Thr Tyr Ala Glu Asn Ile Thr Lys Asn Asn Ser Gln
123 165 170 175
126 Leu Asn Leu Val Val Ala Ile Asn Tyr Ser Gly Lys Tyr Asp Ile Val
127 180 185 190
130 Gln Ala Cys Gln Ser Ile Ala Leu Lys Val Lys Asp Gly Val Ile Gln
131 195 200 205
134 Pro Glu Glu Ile Asn Glu Phe Thr Ile Glu Asn Glu Leu Gly Thr Asn
135 210 215 220
138 Cys Ile Pro Phe Pro His Pro Asp Leu Leu Ile Arg Thr Ser Gly Glu
139 225 230 235 240
142 Leu Arg Val Ser Asn Phe Phe Leu Trp Gln Leu Ala Tyr Thr Glu Leu
143 245 250 255
146 Tyr Phe Ser Glu Thr Leu Trp Pro Asp Phe Gly Glu Asp Glu Leu Leu
147 260 265 270
150 His Ala Leu Asn Thr Phe Gln His Arg Arg Arg Arg Tyr Gly Gly
151 275 280 285
154 <210> SEQ ID NO: 3
155 <211> LENGTH: 1082
156 <212> TYPE: DNA
157 <213> ORGANISM: Calendula officinalis
159 <400> SEQUENCE: 3
160 atgacattat tttccctaata tactcaatta aaccttggtt agctcctaaa ccacactctt 60
162 ccaactctaa tttcttcaac cgcgtgtcac caataacttc ggggataatt cgttcacoga 120
164 ttacaaatat atcaacggtt ttgagcaatg aaaataccaa actgaaaacc aaaaaaagaa 180
166 gaagtagaat taccaggggg tctcgaagaa gaactaatgc caaaacacgt tgcattcata 240
168 atggatggaa accgtcgatg ggcggtggaa aaaggttggg ctccaatgac gggtcatagt 300
170 gccatgagaa agacgcttca atctctcctt tttcgatggt ccaaattcaa aatcaaagcg 360

```

## RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/10/069,425A

TIME: 12:30:18

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\06252003\J069425A.raw

```

172 gtatcgattt atgcattttc taccgaaaat tggactcgcc cgaaggaaga agttgatttc 420
174 ctaatggaga tgtatgaaga tttattgagg acagatgctg aggagctott aagtccttgg 480
176 tgtcgagtaa gcataatggg gaaaaagacc aaccttcga aatcactaca aaagttatgc 540
178 atcgaaatag aagaaaaatc aagagccaat tcaggaaccc atgttaacta tgcactcaac 600
180 tacagtggaa aatacgacat aatcgaagct tgtaaaagcg tcgctacaaa agtcaaggat 660
182 ggtgttatta ttccaaaaca gatcgacgaa aaatatttca aacaagaact cggtagccaa 720
184 atgatcgatt ttccattacc tgacctagtt atacgtacaa gcgggggaaat taggcttagt 780
186 aatttcatgc tatggcagat ggcgtatagc gagctttatt tcacggataa atactttccg 840
188 gattttgggg aaaatgatct tatcgaggct ttacttgcac ttcaaaaagt gcgtaaatgt 900
190 taataacttg ttgtgggtaa gacgagtgtg gtagaatatc aataaatgac tcgtttcggc 960
192 ggcgttgtgt atgccacatt atatgtctta gtgtctatca gaattcgaat ttgatttata 1020
194 gtcgcttgag atatgaaaac ttattatatt tgttcgatca aaaaaaaaaa aaaaaaaaaa 1080
196 aa 1082
199 <210> SEQ ID NO: 4
200 <211> LENGTH: 228
201 <212> TYPE: PRT
202 <213> ORGANISM: Calendula officinalis
204 <400> SEQUENCE: 4
206 Met Pro Lys His Val Ala Phe Ile Met Asp Gly Asn Arg Arg Trp Ala
207 1 5 10 15
210 Val Glu Lys Gly Trp Ser Pro Met Thr Gly His Ser Ala Met Arg Lys
211 20 25 30
214 Thr Leu Gln Ser Leu Leu Phe Arg Cys Ser Lys Phe Lys Ile Lys Ala
215 35 40 45
218 Val Ser Ile Tyr Ala Phe Ser Thr Glu Asn Trp Thr Arg Pro Lys Glu
219 50 55 60
222 Glu Val Asp Phe Leu Met Glu Met Tyr Glu Asp Leu Leu Arg Thr Asp
223 65 70 75 80
226 Ala Glu Glu Leu Leu Ser Leu Gly Cys Arg Val Ser Ile Met Gly Lys
227 85 90 95
230 Lys Thr Asn Leu Pro Lys Ser Leu Gln Lys Leu Cys Ile Glu Ile Glu
231 100 105 110
234 Glu Lys Ser Arg Ala Asn Ser Gly Thr His Val Asn Tyr Ala Leu Asn
235 115 120 125
238 Tyr Ser Gly Lys Tyr Asp Ile Ile Glu Ala Cys Lys Ser Val Ala Thr
239 130 135 140
242 Lys Val Lys Asp Gly Val Ile Ile Pro Lys Gln Ile Asp Glu Lys Tyr
243 145 150 155 160
246 Phe Lys Gln Glu Leu Gly Thr Lys Met Ile Asp Phe Pro Tyr Pro Asp
247 165 170 175
250 Leu Val Ile Arg Thr Ser Gly Glu Ile Arg Leu Ser Asn Phe Met Leu
251 180 185 190
254 Trp Gln Met Ala Tyr Ser Glu Leu Tyr Phe Thr Asp Lys Tyr Phe Pro
255 195 200 205
258 Asp Phe Gly Glu Asn Asp Leu Ile Glu Ala Leu Leu Ala Phe Gln Lys
259 210 215 220
262 Val Arg Lys Cys
263 225
266 <210> SEQ ID NO: 5

```

## RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/10/069,425A

TIME: 12:30:18

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\06252003\J069425A.raw

267 &lt;211&gt; LENGTH: 1071

268 &lt;212&gt; TYPE: DNA

269 &lt;213&gt; ORGANISM: Hevea brasiliensis

271 &lt;400&gt; SEQUENCE: 5

```

272 tctcattcga gtgctcaagt tgcaaaccac ttttgatttt ggaggattta ccgagtcacc      60
274 tacaggcttc gggttaaagc atcgtgatgt gggtttaagg aaatggaatt atataccagt      120
276 taagtcagtg atttaaggaa aatggaatta tacaacgggtg agaggccaag tgtgttcaga      180
278 ctttttagga agtatatgag aaaaggggta tatagcatcc taaccagggg tcccatccct      240
280 actcatattg ccttcatatt ggatggaaac aggagggttg ctaagaagca taaactgcc      300
282 gaaggaggtg gtcataaggc tggattttta gctcttctga acgtactaac ttattgctat      360
284 gagttaggag tgaaatatgc gactatctat gccttttagc tcgataaatt tcgaaggaaa      420
286 cctcatgagg ttcagtacgt aatggatcta atgctggaga agattgaagg gatgatcatg      480
288 gaagaaagta tcatcaatgc atatgatatt tgcgtacgtt ttgtgggtaa cctgaagctt      540
290 ttaagtgagc ccgtcaagac cgcagcagat aagattatga gggctactgc caacaattcc      600
292 aaatgtgtgc ttctcattgc tgtatgctat acttcaactg atgagatcgt gcatgctgtt      660
294 gaagaatcct ctgaattgaa ctccaatgaa gtttgtaaca atcaagaatt ggaggaggca      720
296 aatgcaactg gaagcagtag tgtgattcaa actgagaaca tggagtcgta ttctggaata      780
298 aaacttgtag accttgagaa aaacacctac ataaatcctt atcctgatgt tctgattcga      840
300 acttctgggg agaccgtctc gagcaactac ttactttggc agactactaa ttgcatactg      900
302 tattctcctt atgcactgtg gccagagatt ggtcttcgac acgtggtgtg gtcagtaatt      960
304 aacttccaac gtcattattc ttacttggag aaacataagg aatacttaaa ataatttggt      1020
306 tctgttccta gctcatcctg ccttattccg ataggttaag cttaagcata t              1071

```

309 &lt;210&gt; SEQ ID NO: 6

310 &lt;211&gt; LENGTH: 290

311 &lt;212&gt; TYPE: PRT

312 &lt;213&gt; ORGANISM: Hevea brasiliensis

314 &lt;400&gt; SEQUENCE: 6

```

316 Met Glu Leu Tyr Asn Gly Glu Arg Pro Ser Val Phe Arg Leu Leu Gly
317 1          5          10          15
320 Lys Tyr Met Arg Lys Gly Leu Tyr Ser Ile Leu Thr Gln Gly Pro Ile
321          20          25          30
324 Pro Thr His Ile Ala Phe Ile Leu Asp Gly Asn Arg Arg Phe Ala Lys
325          35          40          45
328 Lys His Lys Leu Pro Glu Gly Gly Gly His Lys Ala Gly Phe Leu Ala
329          50          55          60
332 Leu Leu Asn Val Leu Thr Tyr Cys Tyr Glu Leu Gly Val Lys Tyr Ala
333 65          70          75          80
336 Thr Ile Tyr Ala Phe Ser Ile Asp Asn Phe Arg Arg Lys Pro His Glu
337          85          90          95
340 Val Gln Tyr Val Met Asp Leu Met Leu Glu Lys Ile Glu Gly Met Ile
341          100         105         110
344 Met Glu Glu Ser Ile Ile Asn Ala Tyr Asp Ile Cys Val Arg Phe Val
345          115         120         125
348 Gly Asn Leu Lys Leu Leu Ser Glu Pro Val Lys Thr Ala Ala Asp Lys
349          130         135         140
352 Ile Met Arg Ala Thr Ala Asn Asn Ser Lys Cys Val Leu Leu Ile Ala
353 145         150         155         160
356 Val Cys Tyr Thr Ser Thr Asp Glu Ile Val His Ala Val Glu Glu Ser
357          165         170         175

```

## RAW SEQUENCE LISTING

DATE: 06/25/2003

PATENT APPLICATION: US/10/069,425A

TIME: 12:30:18

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\06252003\J069425A.raw

```

360 Ser Glu Leu Asn Ser Asn Glu Val Cys Asn Asn Gln Glu Leu Glu Glu
361          180          185          190
364 Ala Asn Ala Thr Gly Ser Ser Thr Val Ile Gln Thr Glu Asn Met Glu
365          195          200          205
368 Ser Tyr Ser Gly Ile Lys Leu Val Asp Leu Glu Lys Asn Thr Tyr Ile
369          210          215          220
372 Asn Pro Tyr Pro Asp Val Leu Ile Arg Thr Ser Gly Glu Thr Arg Leu
373 225          230          235          240
376 Ser Asn Tyr Leu Leu Trp Gln Thr Thr Asn Cys Ile Leu Tyr Ser Pro
377          245          250          255
380 Tyr Ala Leu Trp Pro Glu Ile Gly Leu Arg His Val Val Trp Ser Val
381          260          265          270
384 Ile Asn Phe Gln Arg His Tyr Ser Tyr Leu Glu Lys His Lys Glu Tyr
385          275          280          285
388 Leu Lys
389          290
392 <210> SEQ ID NO: 7
393 <211> LENGTH: 1000
394 <212> TYPE: DNA
395 <213> ORGANISM: Hevea brasiliensis
397 <400> SEQUENCE: 7
398 cgggttaagt cagtgattta aggaaaatgg aattatacaa cggtgagagg ccaagtgtgt      60
400 tcagactttt agagaagtat atgagaaaag ggttatatag catcctaacc cagggtccca      120
402 tccctactca tattgccttc atattggatg gaaacaggag gtttgctaag aagcataaac      180
404 tgccagaagg aggtgggtcat aaggctggat ttttagctct tctgaacgta ctaacttatt      240
406 gctatgagtt aggagtgaat tatgcgacta tctatgcctt tagcatcgat aattttcgaa      300
408 ggaaacctca tgaggttcag tacgtaatgg atctaattgt ggagaagatt gaaggatga      360
410 tcatggaaga aagtatcatc aatgcatatg atatttgcgt acgttttgtg ggtaacctga      420
412 agctttttaag tgagccagtc aagaccgcag cagataagat tatgagggct actgccaca      480
414 attccaaatg tgtgcttctc attgctgtat gctatacttc aactgatgag atcgtgcatg      540
416 ctgttgaaga atcctctgaa ttgaactcca atgaagtttg taacaatcaa gaattggagg      600
418 aggcaaatgc aactggaagc agtactgtga ttcaaactga gaacatggag tcgtattctg      660
420 gaataaaact tgtagacctt gagaaaaaca cctacataaa tccttatcct gatgttctga      720
422 ttcgaacttc tggggagacc cgtctgagca actacttact ttggcagact actaattgca      780
424 tactgtattc tccttatgca ctgtggccag agattgggtc tcgacacgtg gtgtgggtcag      840
426 taattaactt ccaacgtcat tattcttact tggagaaaca taaggaatac ttaaaataat      900
428 ttgtttctgt tcctagctca tcctgcctta ttgcgcatag ttaagcttaa gcatatcctt      960
430 gtggaataaa ctcgacact taattaagcc ggtattttgt      1000
433 <210> SEQ ID NO: 8
434 <211> LENGTH: 290
435 <212> TYPE: PRT
436 <213> ORGANISM: Hevea brasiliensis
438 <400> SEQUENCE: 8
440 Met Glu Leu Tyr Asn Gly Glu Arg Pro Ser Val Phe Arg Leu Leu Glu
441 1          5          10          15
444 Lys Tyr Met Arg Lys Gly Leu Tyr Ser Ile Leu Thr Gln Gly Pro Ile
445          20          25          30
448 Pro Thr His Ile Ala Phe Ile Leu Asp Gly Asn Arg Arg Phe Ala Lys
449          35          40          45

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 06/25/2003  
PATENT APPLICATION: US/10/069,425A      TIME: 12:30:19

Input Set : A:\PTO.AMC.txt  
Output Set: N:\CRF4\06252003\J069425A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:21; Xaa Pos. 2,3,8,10  
Seq#:22; Xaa Pos. 3,7,10,12,20,21,23